ABSTRACT OF THE DISCLOSURE

To enhance the distribution or backup of data similar to previously distributed or backed up data, to provide similarity testing, and other benefits, a first body of data T of size n is compressed with respect to a second body of data S of size m *in-place*; that is, the memory containing S is overwritten from left to right so that at no time is more than a total of $MAX\{m,n\} + O(1)$ memory used.